

Automatic Extraction of Concepts from Gastrointestinal Endoscopy Reports

Michele Tringali,[†] Marcelo Fiszman,[‡] Thomas Rindflesch,[‡] Olivier Bodenreider[‡]

[†]Azienda Ospedaliera Santa Maria della Misericordia, Udine, Italy

[‡]U.S. National Library of Medicine, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA

Automatic linguistic tools can be used to extract clinical concepts from free text reports. This structured information can be used for medical decision support, quality assurance, epidemiology studies, and clinical research. Recent efforts have been directed, for example, at automatically processing radiology reports and discharge summaries ([1] and [2]). We report preliminary results from a project to identify and extract useful concepts from free text gastrointestinal (GI) endoscopy reports. Such processing provides a first step toward the semantic annotation of text and images in support of enhanced information retrieval applications.

Methods: A sample of 25 esophagogastroduodenoscopy reports from Clarkson University Hospital, Omaha, Nebraska, USA was randomly selected for analysis. A gold standard was produced by a board certified gastroenterologist (MT), who selected concepts along with their semantic types from the Unified Medical Language System[®] (UMLS)[®] Metathesaurus[®] (2002) to represent the GI content of these reports. MetaMap [3], a program designed to map free text to concepts in the Metathesaurus, was then used to automatically generate concepts (and semantic types) for each report. The output from Metamap was compared to the gold standard, and an exact match was required for a MetaMap concept to be considered correct.

Performance metrics were calculated for concepts with selected UMLS semantic types in the following three clinically relevant groups: Anatomy ('Body Part, Organ, or Organ Component', 'Body Location or Region'); Problem ('Finding', 'Disease or Syndrome', 'Sign or Symptom'); and Procedure ('Diagnostic Procedure', 'Therapeutic or Preventive Procedure').

Results: The overall recall for the three groups combined was 0.64, while precision was 0.62. Results by the groups noted above are given in Table 1.

Group	Recall	Precision
Anatomy	0.61	0.72
Problem	0.70	0.65
Procedure	0.57	0.49

Table 1 – Performance measures

Discussion: The main objective of this preliminary study was to determine baseline

performance of MetaMap in identifying useful GI concepts in endoscopy reports, and the program was thus used without modification or enhancement. The results are nonetheless promising.

The majority of false positives are due to word sense ambiguity, and the majority of false negatives are associated with missing synonyms in the Metathesaurus. As an example of the latter, the text *second portion of the duodenum* does not appear in the Metathesaurus, although "Descending portion of the duodenum", "Descending part of duodenum" and "Second part of duodenum" do occur.

The less accurate performance on procedures, particularly with respect to precision, is largely due to the general phenomenon of ambiguity in the Metathesaurus. For example, text *esophagus* matches to both "Esophagus" and "Procedures of the esophagus," thus generating the latter concept as a false positive. Correcting this problem would increase precision from 0.49 to 0.65 for the Procedure group.

Other optimizations that could increase the performance of the system are possible. Changes could be made to MetaMap to prefer pre-coordinated instead of atomic terms (e.g., for *gastrointestinal bleeding*). Selecting concepts that are specific to the gastrointestinal domain is also possible.

Conclusion: This study is a preliminary attempt to extract concepts from gastrointestinal endoscopy reports. Although it is limited by the modest number of reports and by the use of only one expert in creating the gold standard, the results provide valuable guidance toward achieving high-quality processing of clinical text.

References

- [1] Fiszman M, et al. Automatic detection of acute bacterial pneumonia from chest X-ray reports. J Am Med Inform Assoc. 2000; 7(6):593-604.
- [2] Friedman C. Automating a severity score guideline for community-acquired pneumonia employing medical language processing of discharge summaries. Proc AMIA Symp. 1999;256-60.
- [3] Aronson AR. Effective mapping of biomedical text to UMLS Metathesaurus: The MetaMap program. Proc AMIA Symp. 1999;17-21.